WHAT IS CLAIMED IS:

1. A computer implemented method for modeling ligand-protein binding interactions, comprising:

providing structural information describing the structure of a protein and a set of one or more ligands;

using the structural information for the protein to identify a binding region of the protein;

identifying a plurality of preferred binding conformations for each of the set of ligands in the binding region;

optimizing the preferred binding conformations using annealing molecular dynamics, the annealing molecular dynamics including solvation effects;

calculating a binding energy for each of the set of ligands in the corresponding optimized preferred binding conformations; and

selecting for each of the set of ligands the lowest calculated binding energy in the optimized preferred binding conformations, and outputting the selected calculated binding energies as the predicted binding energies for each of the set of ligands.

2. The method of claim 1, wherein:

the binding region is a known binding region defined by the structural information.

3. The method of claim 1, wherein:

the binding region is an unknown binding region; and

using the structural information for the protein to identify a binding region of the ligand in the protein comprises predicting a probable binding region based at least in part on the structural information.

4. The method of claim 3, wherein predicting a probable binding region comprises:

mapping the empty volumes available for ligand binding in the protein to identify one or more potential binding regions;

generating initial conformations for one or more ligands known to bind the protein using docking techniques in each of the one or more potential binding regions;

selecting from the initial conformations for each of the known ligands a plurality of best conformations in each of the potential binding regions and scoring an energy function for each of the best conformations; and

identifying the probable binding site based on a spatial location of the conformations having the lowest energy scores.

5. The method of claim 4, further comprising:

before scoring the energy function for each of the best conformations, optimizing the selected best conformations to obtain a set of energy-minimized conformations for each of the known ligands in each of the potential binding regions;

wherein the energy function is scored for each of the energy-minimized conformations.

6. The method of claim 4, further comprising:

before scoring the energy function for each of the best conformations, calculating for each of the best conformations a percentage of the ligand surface area buried within the protein for the conformation;

wherein the energy function is scored only for a subset of the best conformations having a calculated percentage of the ligand surface area buried within the protein exceeding a predetermined surface area threshold.

7. The method of claim 1, wherein identifying the preferred binding conformations for each of the set of ligands comprises:

generating initial conformations for each of the set of ligands in the binding region using docking techniques; and

selecting from the initial conformations for each of the ligands a plurality of best conformations.

8. The method of claim 7, further comprising:

after selecting the best conformations, optimizing the selected best conformations to obtain a set of energy-minimized conformations for each of the ligands;

wherein the preferred binding conformations comprise the energy-minimized conformations.

9. The method of claim 1, wherein:

the annealing molecular dynamics includes a full atom force field.

10. The method of claim 1, wherein:

the solvation effects include a continuum description of solvation.

11. The method of claim 1, wherein:

the solvation effects include a surface-area based solvation model.

12. The method of claim 1, wherein:

calculating a binding energy for each of the set of ligands includes taking the difference in the ligand energy in the receptor and in solution.

13. The method of claim 1, wherein:

the binding energy is calculated for a ligand according to a scoring function comprising subtracting the free energy of the ligand in water from the energy of the ligand in the protein.

14. The method of claim 1, wherein:

the binding energy is calculated for a ligand according to a scoring function comprising subtracting the free energy of the protein and the free energy of the ligand from the free energy of the ligand in the protein.

15. The method of claim 1, further comprising:

identifying from the set of ligands one or more ligands predicted to have high binding affinity based on the calculated binding energy of the ligands in the binding site.

16. The method of claim 1, wherein:

the protein is a globular protein or a transmembrane protein.

17. A computer-implemented method for predicting the structure of a protein binding site for a protein having an unknown binding site, the method comprising:

providing structural information describing the structure of a protein having an unknown binding site and a set of one or more ligands known to bind to the protein;

using the structural information for the protein to identify a plurality of potential binding regions of the protein;

generating initial conformations for one or more of the ligands using docking techniques in each of the potential binding regions;

selecting from the initial conformations for each of the ligands a plurality of best conformations in each of the potential binding regions and scoring an energy function for each of the best conformations;

identifying the probable binding site based on a spatial location of the conformations having the lowest energy scores; and

outputting structure information describing the three-dimensional structure of the probable binding site.

18. The method of claim 17, further comprising:

before scoring the energy function for each of the best conformations, optimizing the selected best conformations to obtain a set of energy-minimized conformations for each of the ligands in each of the potential binding regions;

wherein the energy function is scored for each of the energy-minimized conformations.

19. The method of claim 17, further comprising:

before scoring the energy function for each of the best conformations, calculating for each of the best conformations a percentage of the ligand surface area buried within the protein for the conformation;

wherein the energy function is scored only for a subset of the best conformations having a calculated percentage of the ligand surface area buried within the protein exceeding a predetermined surface area threshold.

20. A computer-implemented virtual screening method for screening a ligand library, the method comprising:

receiving protein structural information describing the structure of a protein;

receiving ligand structural information describing the structure of a plurality of ligands in a ligand library;

receiving an input specifying a desired number of candidate ligands to be identified in the ligand library;

using the structural information for the protein to identify a binding region of the protein;

generating a set of initial binding conformations for each of the ligands in the binding region;

calculating an energy function for each of the initial binding conformations and selecting for each of the ligands a plurality of the initial binding conformations having the lowest calculated energy as a set of best conformations;

optimizing the best conformations;

calculating a binding energy for each of the ligands in the corresponding optimized best conformations; and

selecting from the plurality of ligands a set of the desired number of candidate ligands having the lowest calculated binding energy in the optimized best binding conformations, and outputting the selected set of candidate ligands.

21. The method of claim 20, wherein:

the plurality of ligands comprises at least 500 ligands.

22. The method of claim 20, wherein:

the plurality of ligands comprises at least 1,000 ligands.

23. The method of claim 20, wherein:

the plurality of ligands comprises at least 5,000 ligands.

24. The method of claim 20, wherein:

the plurality of ligands comprises at least 10,000 ligands.

25. The method of claim 20, wherein:

the plurality of ligands comprises at least 50,000 ligands.

26. The method of claim 20, wherein:

the plurality of ligands comprises at least 100,000 ligands.

27. The method of claim 20, wherein:

calculating a binding energy for each of the set of ligands includes taking the difference in the ligand energy in the receptor and in solution.

28. The method of claim 20, wherein:

the binding energy is calculated for a ligand according to a scoring function comprising subtracting the free energy of the ligand in water from the energy of the ligand in the protein.

29. A computational model of a ligand-protein complex for a protein having an unknown binding site, the model comprising:

a computer-readable memory storing data describing an optimized preferred binding conformation for the protein and a

ligand known to bind to the protein, the optimized binding conformation being generated according to the method claim 1.

30. A computational model of a predicted structure for a protein binding site for a protein having an unknown binding site, the model comprising:

a computer-readable memory storing data describing the three-dimensional structure of the probable binding site for the protein generated according to the method claim 15.

31. A computer program product on a computer-readable medium for modeling ligand-protein binding interactions, the computer program product comprising instructions operable to cause a programmable processor to:

provide structural information describing the structure of a protein and a set of one or more ligands;

use the structural information for the protein to identify a binding region of the protein;

identify a plurality of preferred binding conformations for each of the set of ligands in the binding region;

optimize the preferred binding conformations using annealing molecular dynamics, the annealing molecular dynamics including solvation effects;

calculate a binding energy for each of the set of ligands in the corresponding optimized preferred binding conformations; and

select for each of the set of ligands the lowest calculated binding energy in the optimized preferred binding conformations, and output the selected calculated binding energies as the predicted binding energies for each of the set of ligands.

32. A computer program product on a computer-readable medium for predicting the structure of a protein binding site for a protein having an unknown binding site, the computer program product comprising instructions operable to cause a programmable processor to:

provide structural information describing the structure of a protein having an unknown binding site and a set of one or more ligands known to bind to the protein;

use the structural information for the protein to identify a plurality of potential binding regions of the protein;

generate initial conformations for one or more of the ligands using docking techniques in each of the potential binding regions;

select from the initial conformations for each of the ligands a plurality of best conformations in each of the

potential binding regions and score an energy function for each of the best conformations;

identify the probable binding site based on at least one of a percentage surface area of the ligand buried in the protein or a spatial location of the conformation having the lowest energy score; and

output structure information describing the three-dimensional structure of the probable binding site.

33. A computer program product on a computer-readable medium for screening a ligand library, the computer program product comprising instructions operable to cause a programmable processor to:

receive protein structural information describing the structure of a protein;

receive ligand structural information describing the structure of a plurality of ligands in a ligand library;

receive an input specifying a desired number of candidate ligands to be identified in the ligand library;

use the structural information for the protein to identify a binding region of the protein;

generate a set of initial binding conformations for each of the ligands in the binding region;

calculate an energy function for each of the initial binding conformations and selecting for each of the ligands a plurality of the initial binding conformations having the lowest calculated energy as a set of best conformations;

optimize the best conformations;

calculate a binding energy that includes solvation for each of the ligands in the corresponding optimized best conformations; and

select from the plurality of ligands a set of the desired number of candidate ligands having the lowest calculated binding energy in the optimized best binding conformations, and output the selected set of candidate ligands.

34. A computer-implemented method of generating a pharmacophore, comprising:

providing structural information describing the structure of a protein and a set of one or more ligands known to bind to the protein;

using the structural information for the protein to identify a binding region of the protein;

identifying a plurality of preferred binding conformations for each of the set of ligands in the binding region;

optimizing the preferred binding conformations using annealing molecular dynamics, the annealing molecular dynamics including solvation effects;

calculating a binding energy for each of the set of ligands in the corresponding optimized preferred binding conformations;

selecting for each of the set of ligands the optimized preferred binding conformation having the lowest calculated binding energy;

generating a pharmacophore model based at least in part on the selected optimized preferred binding conformations, the pharmacophore model defining a pattern of ligand features predicted to be required for binding to the protein; and

outputting data representing the pharmacophore model for use in drug design.

35. The method of claim 34, further comprising:

using the pharmacophore model as a template to search a chemical information database to identify one or more molecules predicted to bind to the protein.